AU9341300

(12) PATENT ABSTRACT (11) Document No. AU-A-41300/93
(19) AUSTRALIAN PATENT OFFICE

(54) Title
A METHOD FOR CONTROLLING CONGESTION IN PACKET SWITCHED TELECOMMUNICATIONS NETWORKS

International Patent Classification(s)
(51)[5] **H04L 012/56**

(21) Application No. : **41300/93** (22) Application Date : **17.06.93**

(30) Priority Data

(31) Number	(32) Date	(33) Country
PL3028	18.06.92	AU AUSTRALIA

(43) Publication Date : **23.12.93**

(71) Applicant(s)
TELSTRA CORPORATION LIMITED

(72) Inventor(s)
CHRISTOPHER JOHN O'NEILL

(74) Attorney or Agent
DAVIES COLLISON CAVE , 1 Little Collins Street, MELBOURNE VIC 3000

(57) Claim

1. A method for controlling congestion in packet switched telecommunications networks, comprising:

establishing a queue to buffer packets for transmission on a telecommunications link;

determining, on the basis of at least the packets placed in said queue, whether a user associated with a packet to be placed in said queue has utilised an allowed capacity for said link; and

discarding said packet if said allowed capacity has been utilised.

Regulation 3.2

AUSTRALIA

Patents Act 1990

COMPLETE SPECIFICATION

FOR A STANDARD PATENT

(ORIGINAL)

Name of Applicant: **TELSTRA CORPORATION LIMITED**, A.C.N. 051 775 556, of 242 Exhibition Street, Melbourne 3000, Victoria, Australia

Actual Inventor: Christopher John O'NEILL

Address for Service: **DAVIES COLLISON CAVE**, Patent Attorneys, of 1 Little Collins Street, Melbourne 3000, Victoria, Australia

Invention Title: A METHOD FOR CONTROLLING CONGESTION IN PACKET SWITCHED TELECOMMUNICATIONS NETWORKS

Details of Associated Provisional Application No: PL3028/92

The following statement is a full description of this invention, including the best method of performing it known to us:

A METHOD FOR CONTROLLING CONGESTION IN PACKET SWITCHED TELECOMMUNICATIONS NETWORKS

The present invention relates to a method for controlling congestion in packet switched telecommunications networks.

Congestion in packet switched telecommunications networks, such as asynchronous time division multiplexing (ATM) networks, occurs when capacity allocations for users are exceeded and telecommunications links between users are overloaded with excess packets. In an overload situation, which may occur at the output ports or links of a packet switch, excess packets are simply discarded on a random basis. Protocols are available to users who wish to recover discarded packets but most of these protocols simply retransmit previously received packets, in addition to the packets which have been discarded. Retransmission of successfully received packets exacerbates the overload condition, is a waste of network capacity and gives rise to a situation known as congestion collapse. A severe form of congestion collapse occurs when the available link capacity is less than the average traffic offered for transmission on the link and is known as self sustaining congestion collapse. For satisfactory network performance, self sustaining congestion collapse must be prevented or confined to only occur as a rare event.

Users of the network need to be encouraged to use protocols, known as adaptive protocols, which prevent overload conditions. Adaptive protocols are able to detect when overload is imminent and reduce the amount of traffic offered on a link. The network therefore needs to be configured so as to encourage fair use of network capacity and discourage excessive use, such as by unnecessary retransmission of packets.

The term "user" is used throughout the present specification to refer to the telecommunications connection or traffic stream established by a call, the party which is the originator of the call and/or the party which is the receiver of the call.

One method for controlling congestion is known as fairness queuing which involves implementing a queue for each user of a link at the output ports of a fast packet switch. The queues are served consecutively and, as described hereinafter, packets are automatically discarded when a user's respective queue is full, which would occur during excessive use by the user of the link. Fairness queuing, however, is expensive to implement in high speed packet switched networks as a separate queue needs to exist for each user of each network link. The requirement for a separate queue is particularly onerous in ATM networks where the queues have to be implemented in hardware.

An object of the present invention is to overcome, at least in part, the problems associated with fairness queuing.

In accordance with the present invention there is provided a method for controlling congestion in packet switched telecommunications networks, comprising:

establishing a queue to buffer packets for transmission on a telecommunications link;

determining, on the basis of at least the packets placed in said queue, whether a user associated with a packet to be placed in said queue has utilised an allowed capacity for said link; and

discarding said packet if said allowed capacity has been utilised.

Preferably the link is an output link of a packet switch of said network.

Preferably said allowed capacity has been utilised if the number of packets associated with said user in said queue is greater than or equal to the number of empty packet locations in said queue.

Preferably said allowed capacity has been utilised if said queue is almost full and the number of packets transmitted on said link associated with said user will exceed the average number of packets placed in said queue per user since the queue was empty if said packet is placed in said queue.

Preferably the method includes determining a fair packet transmission share, and wherein said allowed capacity is determined as having been utilised when said share has been exceeded by said user associated with said packet.

Preferred embodiments of the present invention are hereinafter described, by way of example only, with reference to the accompanying drawings, wherein:

Figure 1 is a schematic diagram illustrating the fairness queuing protocol;

Figure 2 is a schematic diagram of an output port of a fast packet switch using a preferred embodiment of a fairness discarding protocol; and

Figure 3 is a graph of a packet count list maintained for a preferred embodiment of the fairness discarding protocol.

Fast packet switches for telecommunications networks may be implemented using a number of different configurations. The switches may be controlled primarily by software, such as described in "Stochastic Fairness Queuing", P.E. McKenney, Internetworking: Research and Experience, Vol. 2, Pages 113 to 131, 1991, or as in an ATM network which meets B–ISDN standards, the switches may be implemented primarily in hardware, such as described in "The Application of Multistage Interconnection Networks to Fast Packet Switching", R. Palmer, C. O'Neill and E. Tirtaatmadja, Journal of Electrical and Electronics Engineering, Australia, Vol. 8, No. 2, Pages 119 to 129, June 1988. The switches all include processing and data storage capabilities which can be adjusted to alter the packet switching and transmission protocols executed by the switches.

The fairness queuing protocol is implemented at each output port of a fast packet switch to avoid congestion of packets routed by the switch to a link. Packets 10 received by an output port 2, as shown in Figure 1, are first processed by a distributor 4 on the basis of the signalling information contained in the packets 10 so as to allocate each packet to a user and to a respective output queue 6 of the port 2. The queues 6 are "first come first serve" queues and one is provided for each user of the port 2. The queues 6 receive packets from the distributor 4 and output the packets on the output link 8 of the port 2 one at a time on a sequential basis. The queues 6 are polled for the release of a

packet in a round robin manner. The sizes of the queue 6 are finite and the queue 6 will begin to fill if a user begins transmitting on the link 8 at a rate which is excessive. If a user is transmitting at a rate which is greater than the user's fair share of the link's capacity, once the user's queue 6 is full, packets associated with the user will begin to be discarded. The fair share is determined by the size of the respective queue 6 and the rate at which it is served.

Therefore as the queues 6 are implemented on a user basis, the fairness queuing method penalises users who choose to exceed their share of a network link's capacity. This fair loss of packets also minimises the effect of overload conditions on users who do not exceed their fair share of capacity and implement adaptive protocols. Fairness queuing allows users to adopt communications protocols which will enable them to achieve transmission at the highest rate that will still avoid loss of packets from the queues 6. However, as discussed previously, fairness queuing requires a queue 6 to be implemented for each user of each network link which is expensive, prohibitively so for ATM networks.

Accordingly, a new method for relieving congestion on a network link has been developed which is described hereinafter and referred to as fairness discarding.

An output port 12 of a fast packet switch, as shown in Figure 2, is provided with a "first come first serve" queue 14 which receives packets from a discarder 16. Packets served from the queue 14 are outputted from the port 12 on an output link 8. All packets of different users destined for the output link 8, and belonging to a particular priority level, are queued in one queue 14. The discarder 16 receives packets 10 sent to the port 12 and determines whether they should be discarded or placed in the queue 14 on the basis of a record 18 which is maintained of packets transmitted and held in the queue 14. On the basis of the record 18, the discarder 16 is able to discard packets associated with a user when it has determined the user has already received a fair share of the output link's capacity. This provides fair sharing of the link's capacity when overload occurs and motivation for users to adopt protocols which alleviate congestion.

The decision to discard or queue a packet made by the discarder 16 may be a function of any variables that are related to the history of packet transmission by the user associated with the packet and, preferably the history of packet transmission from all users. Variables can include the number of packets from each user that are held in the queue 14, the amount of spare buffer space in the queue 14 or a count of the packets which have been sent by each user over a predetermined period of time. The procedures used by the discarder 16, which dictate the records 18 that are maintained, can be classified according to whether they rely just on information associated with the packets held in the queue 14 at the time a new packet is received by the discarder 16, "basic information procedures", or whether they also rely upon information related to packets that have been transmitted and are no longer held in the queue 14, "additional information procedures". The later type of procedure is preferably used when the queue 14 is small, whereas the larger the size of the queue 14, the more information is available and then the basic information procedures are satisfactory. Additional information procedures are, in most cases, more complex than basic information procedures.

A preferred basic information procedure is if a packet received by the discarder 16 belongs to a user which has fewer packets held in the queue or buffer 14 than the number of spare buffer places in the queue 14 then the packet is buffered, and subsequently transmitted, otherwise it is discarded by the discarder 16. The discarder 16 uses the signalling information contained in the packets 10 to determine which user and connection stream a packet belongs to, and on transmission to the queue 14, the packet record 18 is updated to correctly reflect the numbers of packets in the queue 14 for each user.

If there is only one user using the output link 8 then the user will be able to utilise nearly half of the queue 14 without having any packets discarded by the discarder 16, and this is normally an adequate amount of buffer space. If the user overloads the link 8 then packets are discarded once the queue 14 is half filled.

When the link 8 is overloaded and the queue 14 has only one spare space remaining, the only packet that may be queued is one that belongs to a user who does not

already have any packets in the queue 14. The queue 14 will be nearly full most of the time when the link is overloaded. For example, if the queue 14 can hold 98 packets, there are ten active users, and each user has nine packets in the queue 14, giving a total of ninety packets, then there is only eight spare spaces in the queue 14 and a user is only allowed to put another packet in the queue when it has less than eight packets stored therein. Therefore no active user is presently allowed to queue any more packets. Once the oldest packet has been served from the head of the queue 14, the user associated with the packet now has eight packets queued and the queue 14 has nine spare buffer spaces. Therefore that user is allowed to put another packet in the buffer but none of the other users can. Provided users supply packets to the queue whenever spaces are available for them, packets for each user are sequentially buffered in the queue in a cyclical manner. The average packet transmission rate for each user is proportional to the number of packets that each user has in the queue 4, and as each user effectively has the same number of packets in the queue 14 whenever a packet is transmitted, they each achieve the same packet transmission rate. This occurs even if each user attempts to exceed a fair share of capacity, thus fair sharing of capacity of the overloaded link is enforced.

Assuming again each one of ten users has nine packets in the same 98 packet queue, but instead one user acts to minimise congestion collapse by transmitting packets at a constant rate which is just slower than a fair share of link capacity, then at some stage the packet of the constant rate user would be released by the queue 14 and another packet will also be released without the spare spaces being used. When this occurs there will be ten spare buffer places in the queue 14 and any user can then put another packet in the queue. Assuming one of the users that already has nine packets utilises one of the spare spaces, then that user will have ten packets in the queue. Nine spare buffer spaces then exist and the user who had a packet served after the constant rate user is able to utilise one of the spaces as that user only has eight packets queued. Assuming this user utilises the available space, the queue 14 will have eight users with nine packets queued, one user with ten packets queued, the constant rate user with eight packets queued and eight spare buffer places. The constant rate user is now unable to queue another packet in the queue 14 and its packets will be discarded. This is a worse case situation which can occur but when the constant rate user has a packet served it then has seven packets

in the buffer and there are nine spare buffer spaces. Then even if one of the other user's packets are served before the next packet of the constant rate user arrives, and a packet of the other user's is queued, the constant rate user is always able to queue its next packet and have eight packets queued. The constant rate user is therefore guaranteed eight packets in the queue at all times, which gives rise to a guaranteed throughput of 8/90 of the capacity of the output link 8, 90 being the steady state number of packets queued. As there are ten active users, this represents 8/9 of a fair share of the link capacity 10. A measure of fairness, F, can be defined for such worse case situations where a constant rate user is transmitting on a link with users who transmit at the highest successful rate, and is given by the following equation:

$$F = \frac{n_u \times \text{int}\ [(b+1-n_u)/(n_u+1)]}{b + 1 - \text{int}\ [(b+2)/(n_u+1)]} \qquad (1)$$

where n_u is the number of users, b is the packet length of the queue 14, and int refers to an integer function which truncates any remainders.

Fairness, F, declines with a decreasing buffer size and an increasing number of active users. If the number of active users exceeds half the queue size then fairness is zero, whereas if the buffer size is just greater than double the number of active users, fairness is 0.5. To improve fairness in these situations additional information procedures can be used by the discarder 16 but in practice, the procedures may not be required as sufficiently large queues may be implemented so the queue size nearly always substantially exceeds double the number of active users of the link 8.

A preferred additional information procedure is if the queue 14 is nearly full then packets associated with a user are discarded if the user has had nearly the average number of packets accepted for transmission over the period of time since the queue 14 was last empty. Nearly the average number of packets means if another packet is accepted for transmission then, after accepting the packet, more than the average would have been accepted for the user. The procedure requires the discarder 16 to maintain and update the average number of packets accepted for transmission by continually dividing the number of accepted packets by the number of active users, maintaining a tally of the

packets accepted for transmission for each user, and determining what the average will be if a packet is accepted. The procedure effectively decides how full the queue 14 can be before packets are discarded.

If only one user is using the queue 14 and the link 8, the user will be able to fill the buffer as the user will never exceed the new average number of accepted transmissions when a packet is accepted for queuing in the queue 14. Packets will be discarded when the user overloads the link 8 and fills the queue 14.

When the link 8 is overloaded by a number of users, packets will only be accepted which belong to users or streams that are transmitting at less than the average packet rate per user. In the overload condition, the queue 14 will be nearly full most of the time.

The additional information procedure may be implemented when a certain number of spaces in the queue have been used, as the procedure is primarily required when overload is imminent. For example, if the queue 14 holds ten packets and there are ten active users, then the procedure may be implemented only after five packets have been queued. At any instant each user will be able to have some number of packets accepted for queuing and transmission. Users who have fewer than the average number of packets accepted will be able to use any of the spaces in the queue 14 but users who have had more than the average number of packets accepted can only use the first five available spaces before the procedure is used by the discarder 16.

In a situation where one user sends packets at a constant rate and all the other users queue packets as fast as possible, the user sending at the rate which is just less than its fair share, i.e. less than the average number of packets transmitted, will have all packets accepted and transmitted. The other users will share the remaining capacity, but there is not guarantee the remaining capacity will be shared fairly. The other users are only guaranteed to obtain the average accepted number of transmissions for all users. Nevertheless, users will be sufficiently motivated to implement adaptive protocols that control congestion.

To achieve optimum fairness, a fair packet discarding procedure can be used, which involves the discarder 16 keeping a list of the number of packets belonging to each stream that have been accepted and arrived at the queue 14 since it was last empty. The list is arranged in order of numbers of packets for each stream. When packets were added to the queue 14 or released therefrom, the list is updated, if necessary. The procedure determines which packet streams have sent more packets to the link 8 than a fair share. Packets received by the discarder 16 that belong to streams which have exceeded their fair share are discarded. The procedure, as described hereinafter, determines what constitutes a fair share on the basis of the maintained list of packets.

For example, Figure 3 illustrates a graph 24 of the number of packets which have been received on an output port 2 for eight streams since the queue 14 was last empty. The vertical axis 19 represents the number of packets received by the port 2 and the horizontal axis 21 represents the number allocated to each of the streams. The graph 24 represents the information maintained in the record 18. Assuming there are six packets queued, this means that six packets in excess of the transmission capacity of the link have arrived since the last time the queue was empty and have therefore been queued. The excess, bounded by the dotted lines 20 and 22 in the graph 24, may be considered to have been produced by streams 1, 2, 3 and 4, 2.75 from stream 1, 1.75 from stream 2, and 0.75 from each of the streams 3 and 4. The fair share, as indicated by the dotted line 26, for each stream can be arbitrarily determined to be that which, if exceeded, would cause more than four packets to be queued. In other words, if the streams had only transmitted their fair share of capacity then there would only be four packets queued. Therefore streams 1 and 2 have exceeded their fair share by 1.5 and 0.5 packets respectively, and are marked so that any new packets received from the streams are discarded. Packets from streams 1 and 2 are discarded until the fair share value increases so as to be above the number of packets that have been received and accepted and queued for transmission for the streams.

The fair share value is incrementally calculated when a new packet is added to the queue 14 or when a packet is served from the queue. The new fair share value S_f(new) is advantageously calculated from the old fair share value S_f(old), as described

hereinafter, and the values may be real values, not just integers.

When a new packet is added to the queue 14, the maintained count of packets that have been accepted on its stream is increased by one and:

$$S_f(\text{new}) = S_f(\text{old}) - 1/N_x \tag{2}$$

if $N_x > 0$ where N_x is the number of streams previously exceeding the fair share value. If $N_x = 0$:

$$S_f(\text{new}) = S_f(\text{old}) \tag{3}$$

The above is valid provided $S_f(\text{new})$ is greater than S_{1n}, which is the number of accepted packets for the stream which has the greatest number of packets after the streams of the graph 24 that exceed $S_f(\text{old})$. If this is not the case, then the new fair share value is given by:

$$S_f(\text{new}) = S_{1n} - (1 - (S_f(\text{old})-S_{1n})N_x)/(N_x+1) \tag{4}$$

Again, equation 4 is only valid as long as $S_f(\text{new})$ is greater than S_{2n} which is the number of packets in the stream which has the greatest number of packets after the stream with S_{1n} packets. If equation 4 is not valid, then the new fair share is given by:

$$S_f(\text{new}) = S_{2n} - (1 - (S_f(\text{old})-S_{1n})N_x - (S_{1n}-S_{2n})(N_x+1))/(N_x+2) \tag{5}$$

Similarly, equation 5 is valid only if $S_f(\text{new})$ is greater than S_{3n}, being the number of packets in the stream with the next greatest number of packets after the stream with S_{2n} packets. If equation 5 is not valid, then the next equation for $S_f(\text{new})$ can be determined recursively from equations 2 to 5. A stream is marked to have additional packets which arrive discarded when the number of packets it has already had accepted for transmission is greater than $S_f(\text{new})$.

As discussed above, the maintained count of packets of a stream is increased by one when a packet for the stream is added to the queue 14, and the record of counts of packets is reordered if the new count now exceeds a previously smaller count of packets for another stream. If the new count exceeds $S_f(\text{oid})$ then equations 2 to 5 do not apply

and the new fair share value is given by:

$$S_f(\text{new}) = S_f(\text{old}) - (S_f(\text{old}) - S_{1n})/(N_x+1) \tag{6}$$

The right hand side of equation 6, after some manipulation, is the same as the right hand side of equation 4 except for the term $-1/(N_x+1)$. Therefore equation 6 can be replaced by equation 4 when the newly received packet is added to the count S_{1n} before using equation 4 to determine $S_f(\text{new})$.

When a packet is removed from the queue 14 and placed on the output link 8, the new fair share value is given by:

$$S_f(\text{new}) = S_f(\text{old}) + 1/\max(N_x,1) \tag{7}$$

where max( ) is the maximum of the arguments. Equation 7 is valid provided $S_f(\text{new})$ is less than S_{1g} which is the number of packets on the stream having the least number of packets that exceed $S_f(\text{old})$. If equation 7 is not valid then:

$$S_f(\text{new}) = S_{1g} + (1 - (S_{1g} - S_f(\text{old}))N_x)/(N_x-1) \tag{8}$$

Equation 8 is only valid if $S_f(\text{new})$ is less than S_{2g}, which is the number of packets on the stream having the least number of packets greater than S_{1g}. If equation 8 is not valid then:

$$S_f(\text{new}) = S_{2g} + (1 - (S_{1g} - S_f(\text{old}))N_x - (S_{2g} - S_{1g})(N_x-1))/(N_X-2) \tag{9}$$

Similarly, equation 9 is only valid if $S_f(\text{new})$ is less than S_{3g}, which is the number of packets on the stream having the least number of packets greater than S_{2g}. When equation 9 is not valid, the next equation for $S_f(\text{new})$ can be developed recursively from equations 7 to 9.

Equations 2 to 6 decrease the fair share value when a packet is added to the queue 14 and ensure the value is decreased in an inverse proportion with respect to the number of streams which will now exceed or equal the fair share value, i.e. the decrease is shared amongst those streams. Similarly, equations 7 to 9 increase the fair share value when a packet is served from the queue 14 by such an amount so that the increase is in an

inverse proportion with respect to the number of streams that now exceed the fair share value, i.e. the increase is shared amongst those streams.

The count of packets for each stream is continually maintained and incremented, even when packets are removed from the queue 14, and the counts are only reset to zero when the queue 14 becomes empty. The initial value for the fair share value is preferably somewhere near the size of the queue 14 so as to allow for large bursts of traffic without unnecessarily discarding packets which do not give rise to a high risk of overflow from the queue 14. For example, if the buffer size is seven packets, the initial value for the fair share may be set at five packets.

If desired, depending on traffic requirements, the discarder can be configured so traffic is shared unevenly by users. This can be achieved by introducing selected factors into variables of the procedures used by the discarder which are associated with each user.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method for controlling congestion in packet switched telecommunications networks, comprising:

establishing a queue to buffer packets for transmission on a telecommunications link;

determining, on the basis of at least the packets placed in said queue, whether a user associated with a packet to be placed in said queue has utilised an allowed capacity for said link; and

discarding said packet if said allowed capacity has been utilised.

2. A method as claimed in claim 1, wherein the link is an output link of a packet switch of said network..

3. A method as claimed in claim 1 or 2, wherein said basis of said determining step includes the number of empty packet locations in said queue.

4. A method as claimed in any one of the preceding claims, wherein said basis of said determining step includes the number of packets transmitted on said link.

5. A method as claimed in claim 1, 2 or 3, wherein said basis of said determining step includes the number of packets transmitted on said link by each user of said link.

6. A method as claimed in any one of the preceding claims, wherein said basis of said determining step includes the number of packets each user of said link has in said queue.

7. A method as claimed in claim 1 or 2, wherein said determining step determines said allowed capacity has been utilised if the number of packets associated with said user in said queue is greater than or equal to the number of empty packet locations in said queue.

8. A method as claimed in any one of the preceding claims, wherein said determining step determines said allowed capacity has been utilised if said queue is almost full and the number of packets transmitted on said link associated with said user will exceed the average number of packets placed in said queue per user since the queue was empty if said packet is placed in said queue.

9. A method as claimed in claim 8, wherein said average is the number of packets placed in said queue since said queue was empty divided by the number of users of said link.

10. A method as claimed in any one of the preceding claims, including determining a fair packet transmission share, and wherein said allowed capacity is determined as having been utilised when said fair packet transmission share has been exceeded by said user associated with said packet.

11. A method as claimed in any one of claims 1 to 9, including:

maintaining a record of the number of packets each user of said link has placed in said queue;

determining a number of packets comprising a fair packet transmission share, based on said record; and

determining said allowed capacity as having been exceeded when said user exceeds said fair transmission share.

12. A method as claimed in claim 11, wherein said fair transmission share is adjusted when a packet is placed in said queue and when a packet is served from said queue.

13. A method as claimed in claim 11, wherein said fair transmission share is decreased when a packet is placed in said queue and increased when a packet is served from said queue.

14. A method as claimed in any one of claims 10 to 13, wherein said fair transmission share is set so that if exceeded then more than a predetermined number of packets would

be held in said queue.

15. A method for controlling congestion in packet switched telecommunications networks substantially as hereinbefore described with reference to the accompanying drawings.

16. The steps, features, compositions and compounds disclosed herein or referred to or indicated in the specification and/or claims of this application, individually or collectively, and any and all combinations or any two or more of said steps or features.

DATED this 17th day of June, 1993

TELSTRA CORPORATION LIMITED
By its Patent Attorneys
DAVIES COLLISON CAVE

ABSTRACT

A method for controlling congestion in packet switched telecommunications networks, comprising establishing a queue (14) to buffer packets for transmission on a telecommunications link (8), determining, on the basis of at least the packets placed in the queue (14), whether a user associated with a packet to be placed in the queue (14) has utilised an allowed capacity for the link (8), and discarding the packet if the allowed capacity has been utilised.

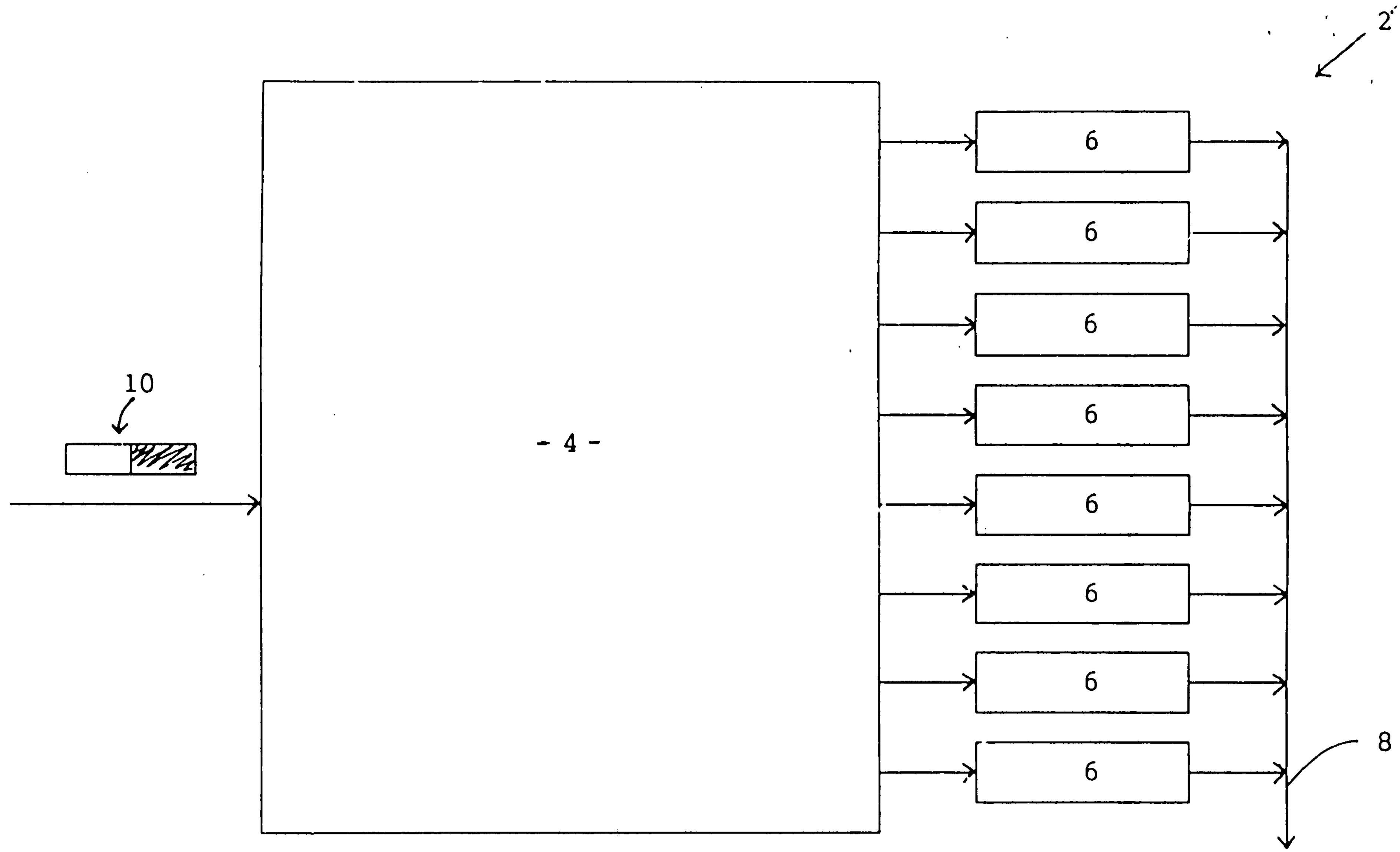

FIGURE 1

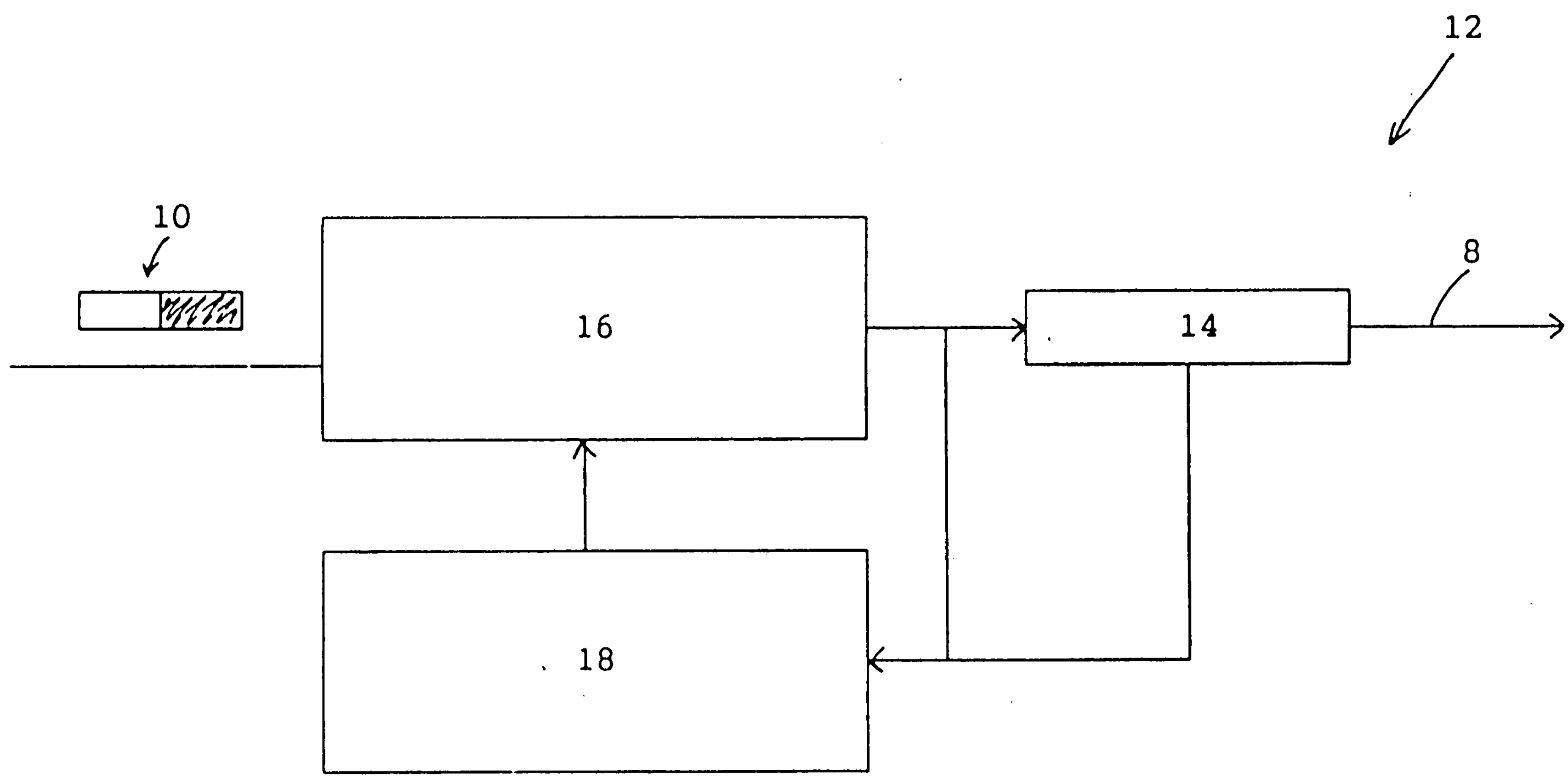

FIGURE 2